Project Plan

# Proposal Role Distribution

# Structure for the Project

This structure was derived based on course expectations and previously completed projects and we are required to find a way to implement this structure into our project.

## Proposal and Task planning (Ross)

* 1. [DONE] Choose a direction of the project
  2. [DONE] Analyse previous projects and understand the requirements
  3. [DONE] Create a rough project plan
  4. [DONE] Assign roles to each group member along with deadlines
  5. [DONE] Find datasets to start with
  6. [DONE] Set up the repository
  7. [DONE] Create an initial draft of project proposal.
  8. [DONE] Edit and refine the proposal
  9. [DONE] Verify that proposal is completed and submit

## Find additional datasets and Source Analysis (Pat)

* 1. [AWAITING] Analyse datasets that are inside the data folder for: (Aim for a point and quick justification for each of the above for every dataset)
     1. Accuracy
     2. Relevance
     3. Reliability
     4. Bias
     5. Use policies and copyright statement of the dataset (can we use it for our project?)
  2. [AWAITING] Determine if there is redundancy of data(data repeats) and if so fix it.

## Data Exploration (Pat, Rick, Ross, Yash)

## Data Preparation and merging (Ross)

## Classification Model, evaluation and tuning

* Make use of various classification/prediction/clustering techniques from the unit
* Evaluate the classification model
  + Use of accuracy
  + F-score
  + Area Under the ROC curve (AUC)
  + Confusion Matrix

## Regression Model, evaluation and tuning (Ross) -30min

* Create model
* Evaluate model #2.2
  + Mean Square Error (MSE)
  + Mean Absolute Error (MAE)
  + R-squared (r2)
  + (MAPE)
* Conclusions based on model

## Intentional Visualisation (Yash, Rick, Pat)

* Determine best ways to visualise results
* Visualise results in succinct, clear and conclusive way

## Presentation Role allocation (Pat, Ross, Yash , Rick) -2hrs

Goal: each person to create 3.75-minute max long video. This is around 600words.

Execution plan: each presentation is to be done once everyone is done with their work.

* Rick
  + Title slide: Project title & group members (names and student ID)
  + Problem statement: What problem are your trying to solve?
* Pat
  + Data sources: What are the data sources you worked on in this project?
* Ross
  + Methods used: What are different methods you applied on chosen datasets?
    - Preparation
    - Data modelling
    - Data evaluation
  + Results of modelling
* Yash/Pat – attempt to contact and no response was provided back so last parts were allocated to them
  + Exploratory analysis results
  + Conclusions: What are your conclusions from this study?